Top charge studies on Monte Carlo

Emanuela Barberis, Iacovos Pilinas, Dennis Sphakov, and Darien Wood.
Northeastern University, Boston, MA 02115.

D∅ Note: ******

Dated: March 25 2004

Abstract

This D∅note describes a technique for measuring the charge of the top quark. According to the Standard Model, the top quark decays into a b-quark and a W-boson. However, the decay of t → b + W- does not violate any laws and this decay gives a charge of -4/3 for the top quark; hence, the motivation of this analysis. This analysis is done on generated ttbar→leptons+jets events and in order to distinguish the two charges, we take advantage of the following tools: kinematic assignment, the sign of the SLT μ and the sign of the isolated μ.

Contents

I. Introduction

The Standard Model (SM) assigns a sign of 2/3 (-2/3) for the top (anti-top) quark. The charge of the top quark can be obtained from the following decay: t→b + W+. On the other hand, one can also observe that the decay t->b + W- is also valid and obeys all the conservation laws. More interesting, the last decay gives a sign of -4/3 (4/3) for the top (anti-top) quark.

In order to distinguish the two cases, we take advantage of the kinematic fitting (HitFit), Soft Lepton Tagging (SLT) and the sign of the isolated muon. In this note, we first describe our method for calculating the required number of data events (with a confidence level 95%) and then we explain our technique for distinguishing the two cases for the charge of the top quark.

For this analysis we use Monte Carlo events for both the signal and background events. For signal events we use the dataset definition: *alpgen_pythia_ttbar-l+jets_175GeV_0.8mb_p14.05.01tmb* and for background we only use W+jjj, W+cjjj, W+ccjj and W+bbjj (or W+4 jets). All MC are top p14 MC and are top analyzed with the latest version of Nefertiti. Also, we only study the muon channel of ttbar→lepton+jets; however, this method should also be done for the electron channel without any major changes.

II. Signal Studies

1. Introduction

We study ttbar→mu+jets Monte Carlo to calculate the number of events that are required to give us a confidence level (CL) of 95% for our measurement. For the purposes of this study, we make use of SLT and kinematic fitting. In the following sections we explain how each technique is used and we also estimate the number of events that are needed (CL of 95%).

2. Pre-selection cuts

The MC events that we use are ttbar→lepton+jets and as we stated above, this analysis is done only on the muon channel. We select events that pass the following criteria[1]:
* General Requirements:
    - Track multiplicity ≥ 3. This is done in an effort to select events with a good quality of the primary vertex requirement.
    - The z position of the primary vertex is required to be within |PVz| < 60cm so that the vertex is inside the fiducial region of the SMT.

---

[1] The preselection cuts are exactly the same as in Ref [1].

- Transverse Missing Energy (MET) has to be greater than 17 GeV
-$\Delta\varphi$ (high pt μ, MET) > 1.2 - (1.2*MET)/3.8
-$\Delta\varphi$ (high pt μ, MET) < 1.3 - (π -1.3*MET)/24
-$\Delta\varphi$ (leading jet, Missing Et) < 2.2 + (π -2.2)*Et/26

- Topological Cuts:
  - The sum of the transverse jet energies to be greater than 110 GeV (Ht>110GeV)
  - Aplanarity to be greater than 0.04.

- Reconstructed Isolated Muon (Reco):
  - Muon has to be isolated.
  - Medium quality muon.
  - nseg=3 and matched to a central track. If a muon candidate has nseg>0 and is also matched to a central track, then the kinematic quantities of this muon are those of the central track and not the muon system.
  - DCA significance of the track with respect to the primary vertex is required to be less than 3.
  - |eta|<2
  - Pt > 20 GeV
  - $\Delta R(\mu_{isolated},$ closest Jet) > 0.5 . This cut translates to the requirement that the isolated muon has to be at least 0.5 radians away from the closest jets.
  - If the muon candidate is cosmic then we reject it.
  - If an event has more than one muon that passes all the cuts for an isolated muon, then we reject this event.

- Generated Muon (MC):
  - Generated isolated muon to come from a W-boson
  - $\Delta R(\mu_{MC}, \mu_{RECO})$ < 0.01 radians
  - Not going through the D$\varnothing$ hole.

The last set of cuts, Generated Muon, is only valid for the Signal Studies. For the Signal and Background studies (section III) we remove the cuts for MC muons since we are treating MC events as data events.

3. Soft Lepton Tagging (SLT) and P_T^{Rel}

i. Introduction

One of the tools that we use to measure the charge of the top quarks is SLT. In addition to the preselection cuts we require at least one b-jet to contain a reconstructed muon. Then we use the sign of the reconstructed muon and assume that is the same as the b-quark that finally decayed into this muon. Obviously, this is not always correct because

of cascade muons (b➔c➔muon) and b oscillations. Consequently, we need to find a quantity that discriminates:

$$b \rightarrow mu+X \quad and \quad b \rightarrow c \rightarrow mu+X.$$

One quantity that does this job is the relative transverse momentum of the muon with respect to its parent jet, also known as P_T^{Rel}. In this section we will give the selection cuts for SLT and will also define the P_T^{Rel} that is used for this analysis.

ii. Soft muon Tagging Requirements

First we match b-quarks to jets with the following criteria:

- b-quarks come from top quarks
- ΔR (b-quark, jet) < 0.5 radians

The purpose of the second cut is to match a b-quark to a jet and this is done if the b-quark direction is within 0.5 radians of a reconstructed jet. If there are more jets that satisfy this requirement we choose the closest jet to the b-quark direction. We are able to apply both cuts because we are using MC events. Of course, later on when we will treat MC events as data events these requirements will not be applied.

If there is at least one b-jet, then we look for a reconstructed muon that obeys the following requirements [1]:

- Reconstructed muon (Reco)
  - Muon is not isolated
  - Medium quality muon
  - |nseg|=3
  - pt > 4 GeV
  - |eta|<2
  - Not a cosmic muon
  - ΔR(muon_reco, b-jet) < 0.5

- Generated muon (MC)
  - Generated muon to come from a b-quark
  - $\Delta R(\mu_{MC}, \mu_{RECO}) < 0.01$ rad
  - Not going through the D∅ hole.

Then we continue by assuming that the sign of the b-quark that produced the b-jet, is the same as the sign of the muon that belongs to this b-jet. As we stated above this is not always correct because of cascade muons (b➔c➔mu+) and b oscillations. We are able to distinguish cascade muons from "good" muons (b➔mu-) by taking advantage of the P_T^{Rel} definition that is used by the b-ID group. This quantity is defined schematically in fig.1 and is calculated from the following expression:
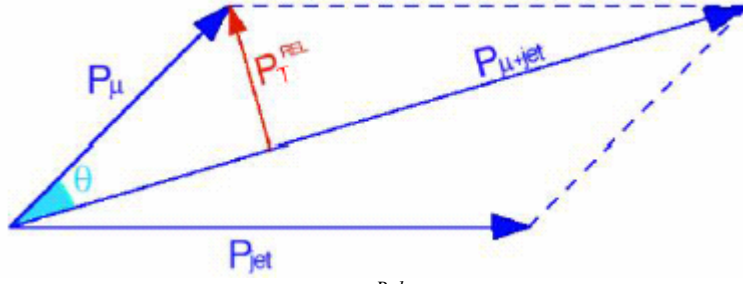
$$P_T{}^{Rel} = P_\mu * \sin(\theta).$$

Figure 1. The definition of $P_T{}^{Rel}$ that is used in this analysis.

The quantity $P_T{}^{Rel}$ has the ability to discriminate muons from b decays because B mesons are heavier than all other sources of muons; thus, muons from b→mu + X have larger momenta relative to their jets axis than any other source of muons. This statement becomes more obvious if we take a look at fig. 2 where we compare $P_T{}^{Rel}$ for b→mu+X and b→c→X+mu.

Figure 2. Comparison of $P_T{}^{Rel}$ for b→μ+X and b→c→μ+X.

Unfortunately, $P_T{}^{Rel}$ takes care only of the cascade muons and not the b oscillations. Also, it is clear from fig.2 that we cannot make a large cut on $P_T{}^{Rel}$ because we will also lose many "good" muons.

Let's define C_{sign}, the percentage of getting the correct sign for b-quark based on the SLT muon and see how this variable behaves with different $P_T{}^{Rel}$ cuts. From fig.3, one sees that $P_T{}^{Rel}$ controls the sign between the b-quark and the SLT muon and with the appropriate cut on $P_T{}^{Rel}$ we can improve the probability of getting the correct sign of the parent b-quark.
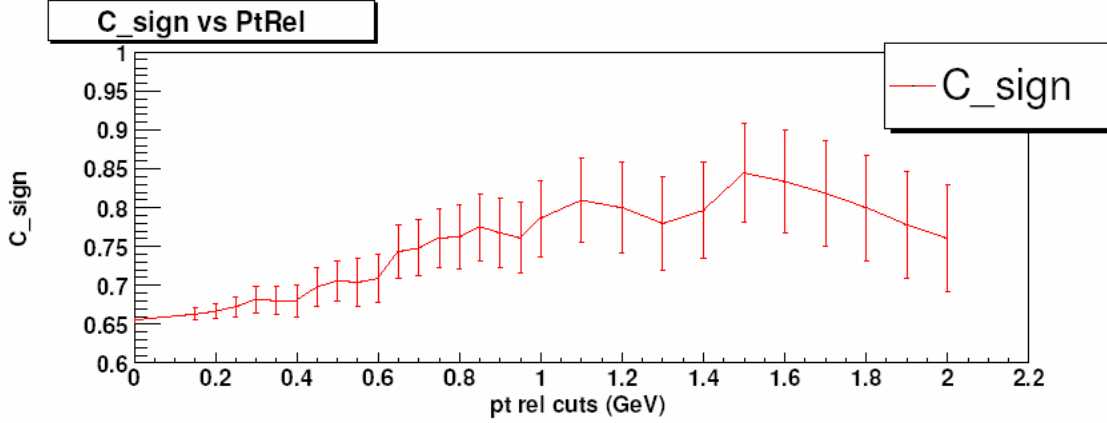
Figure 3. Percentage of getting the correct sign versus P_T^{Rel} cuts. The percentage of getting the correct sign is defined to be:
C_{sign} = correct_sign/(Ncorrect_sign+Nwrong_sign)

4. Assignment of b quarks to top quarks

i. Introduction

Another important piece of information that we need to investigate for measuring the charge of the top quark is the assignment of bottom quarks to top quarks. According to the Standard Model (SM), top quark decays into a b quark and W boson:

$$t \rightarrow b + W+ \qquad \text{and} \qquad tbar \rightarrow bbar + W-$$

On the other hand nothing prevents us from having the following top decays:

$$t \rightarrow b + W- \qquad \text{and} \qquad tbar \rightarrow bbar + W+$$

If the latter decays are correct then the top (anti-top) charge is -4/3 (4/3) and not 2/3 (-2/3). Apparently, we need to investigate the assignment of b-quarks and combine it with the sign studies. The tool that is used for the assignment studies is the HitFit package.

HitFit is used for the mass measurement of the top quark in the lepton+jets channel. In this channel there are at least four jets (see figure 4), which are: two from the hadronic W (W^h), one from the hadronic b quark (b^h) and one from the leptonic b (b^l).
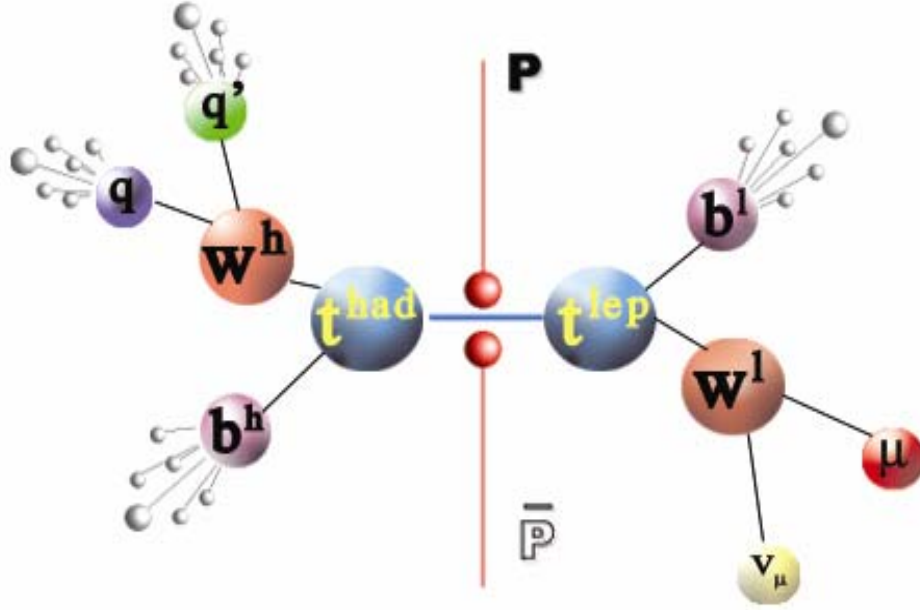

Figure 4. The muon+jets channel.

This algorithm tries to match each parton (bLep, bHad, q, q'), within a simulated event of ttbar→lepton+jets (MC), to the four leading jets according to their P_T (only the four highest P_T jets are used in the fit [2]). If the parton-to-jet matching is successful, then we are able to decode each permutation and see which permutation is assigned to each χ^2. Note that HitFit considers all the possible combinations because we do not know which jet came from which parton. In principle we have 4! permutations for the assignment of the four leading jets to the four partons. However, the assignment and the invariant mass of the W^h (hadronic W) are the same if the two jets assigned to W^h are exchanged. Therefore, there are 4!/2 or 12 permutations.

According to the assignment of b-quarks there are three different categories for the charge measurement. These categories are:

- Correct category:
    Perm 1 : t^{lep}→ b^l + W^l and t^{had}→b^h + W^h / W^h→q+q'
    Perm 2 : t^{lep}→ b^l + W^l and t^{had}→q+ W^h / W^h→ b^h +q'
    Perm 3 : t^{lep}→ b^l + W^l and t^{had}→q'+ W^h / W^h→ q+b^h

- Wrong category:

$$\text{Perm 4}: \ t^{lep} \to b^h + W^l \quad \text{and} \quad t^{had} \to b^l + W^h \ / \ W^h \to q+q'$$
$$\text{Perm 5}: \ t^{lep} \to b^h + W^l \quad \text{and} \quad t^{had} \to q+ W^h \ / \ W^h \to b^l +q'$$
$$\text{Perm 6}: \ t^{lep} \to b^h + W^l \quad \text{and} \quad t^{had} \to q'+ W^h \ / \ W^h \to q+ b^l$$

- Semi-correct category:

$$\text{Perm 7}: \ t^{lep} \to q + W^l \quad \text{and} \quad t^{had} \to b^h + W^h \ / \ W^h \to b^l +q'$$
$$\text{Perm 8}: \ t^{lep} \to q + W^l \quad \text{and} \quad t^{had} \to b^l + W^h \ / \ W^h \to b^h +q'$$
$$\text{Perm 9}: \ t^{lep} \to q + W^l \quad \text{and} \quad t^{had} \to q'+ W^h \ / \ W^h \to b^l +b^h$$
$$\text{Perm 10}: \ t^{lep} \to q' + W^l \quad \text{and} \quad t^{had} \to b^h + W^h \ / \ W^h \to q+b^l$$
$$\text{Perm 11}: \ t^{lep} \to q' + W^l \quad \text{and} \quad t^{had} \to b^l + W^h \ / \ W^h \to q+b^h$$
$$\text{Perm 12}: \ t^{lep} \to q' + W^l \quad \text{and} \quad t^{had} \to q+ W^h \ / \ W^h \to b^l +b^h$$

In this analysis we select the permutation with the lowest χ^2. If the permutation with the lowest χ^2 is in the correct (wrong) category then we know that both b-quarks are assigned to the correct (wrong) top quark. On the other hand, if the selected permutation belongs to the semi-correct category, we need to make sure which b-quark (bLep or bHad) is b-tagged (see II.2) and decide accordingly. For example, permutation 7,8 and 9 give the correct assignment for the hadronic b and the wrong permutation for the leptonic b.

In the previous paragraphs we use the phrase "b is assigned to the correct (or wrong) top" and this statement might cause confusion. We are allowed to use this terminology because the signal studies are performed on simulated events of ttbar→lepton+jets, which are generated based on the SM. Because of that we do not expect to see any decays of the form t→b + W-; however, this might be the case for real data.

Finally, for the assignment studies we see that we have six permutations that give the correct assignment for b quarks and six that assign them incorrectly. These incorrect permutations can be seen as another source of background.

iii. Using b-tagging within HitFit

In an effort to reduce the background due to incorrect assignments we take advantage of b-tagging within HitFit. HitFit can be set to use Soft Lepton Tagging (SLT), Secondary Vertex Tagging (SVT) and Impact Parameter (IP) tagging. In the case of 4 jets with one b-tagged jet, HitFit outputs 6 rather than 12 permutations and in the case of 2 b-tags it outputs only two permutations. A priori, we expect that b-tagging within HitFit will reduce the background caused by incorrect assignments (see figure 5).

With the use of b-tagging we are left with only six permutations (maximum); three of them are correct and three of them are incorrect. Based on which b is b-tagged (bLep or bHad), HitFit outputs the following six permutations (only one b-tagged):

- bLep is b-tagged:
    - Correct Assignment:
        Perm 1 : $t^{lep} \rightarrow b^l + W^l$ and $t^{had} \rightarrow b^h + W^h$ / $W^h \rightarrow q+q'$
        Perm 2 : $t^{lep} \rightarrow b^l + W^l$ and $t^{had} \rightarrow q+ W^h$ / $W^h \rightarrow b^h +q'$
        Perm 3 : $t^{lep} \rightarrow b^l + W^l$ and $t^{had} \rightarrow q'+ W^h$ / $W^h \rightarrow q+b^h$
    - Wrong Assignment:
        Perm 4 : $t^{lep} \rightarrow b^h + W^l$ and $t^{had} \rightarrow b^l + W^h$ / $W^h \rightarrow q+q'$
        Perm 8 : $t^{lep} \rightarrow q + W^l$ and $t^{had} \rightarrow b^l + W^h$ / $W^h \rightarrow b^h +q'$
        Perm 11 : $t^{lep} \rightarrow q' + W^l$ and $t^{had} \rightarrow b^l + W^h$ / $W^h \rightarrow q+b^h$

- bHad is b-tagged:
    - Correct Assignment:
        Perm 1 : $t^{lep} \rightarrow b^l + W^l$ and $t^{had} \rightarrow b^h + W^h$ / $W^h \rightarrow q+q'$
        Perm 7 : $t^{lep} \rightarrow q + W^l$ and $t^{had} \rightarrow b^h + W^h$ / $W^h \rightarrow b^l +q'$
        Perm 10 : $t^{lep} \rightarrow q' + W^l$ and $t^{had} \rightarrow b^h + W^h$ / $W^h \rightarrow q+b^l$
    - Wrong Assignment:
        Perm 4 : $t^{lep} \rightarrow b^h + W^l$ and $t^{had} \rightarrow b^l + W^h$ / $W^h \rightarrow q+q'$
        Perm 5 : $t^{lep} \rightarrow b^h + W^l$ and $t^{had} \rightarrow q+ W^h$ / $W^h \rightarrow b^l +q'$
        Perm 6 : $t^{lep} \rightarrow b^h + W^l$ and $t^{had} \rightarrow q'+ W^h$ / $W^h \rightarrow q+ b^l$

In the case of two b-tagged jets HitFit outputs only two permutations, which are Perm 1 (correct assignment) and Perm 4 (bLep and bHad flipped).

An interesting point from the above categories is that we have only ten permutations in total. This is because Perm 9 and Perm 12 are not used as far as we have at least one b-tagged jet. In these two permutations HitFit assigns both b-quarks to the hadronic W. The kinematic fitting package is smart enough to not use these permutations for the reason that a b-tagged jet will only have to be assigned to tLep or tHad. However; we see that the total number of permutations in the above categories, bLep is tagged or bHad is tagged, is twelve. This is because the two categories share Perm 1 and Perm 9.

iv. Additional Cuts

Until this point we have introduced the preselection cuts (section II.2) and P_T^{Rel} cut (section II. 3). As one can realize from reading the section describing how the kinematic fitting package is used, we have to use two additional cuts:

- χ^2 of the fitting for the lowest χ^2 to be positive
- Parton-to-jet matching to be successful.

The first requirement uses the output of HitFit (TopFitArray_chisq[1]) and tells us how good the fit is. When we use b-tagging within HitFit we have to use TopFitArray_bchisq[1], which has the same meaning as TopFitArray_chisq[1]. The output of HitFit that has a b in front of all the arrays is the output using b-tagging within HitFit. Also, we have to make sure that TopFitArray_chisq and TopFitArray_bchisq are different (according to the order of the χ^2) because if there is no b-tagged jet then the two arrays are exactly the same.

The second requirement is important when we decode each permutation. All four partons (bLep, bHad, q1 and q2) have to be matched to a jet with an index of less or equal to 3 (starting to count from0) for the reason that HitFit uses only the four leading jets.

iv. Setting the top mass=175GeV within HitFit

Since we are not concerned with measuring the mass of the top quark, we are capable of running HitFit with top mass=175 GeV (the simulated events are generated with top mass=175GeV). Again we expect to reduce the background due to incorrect assignments because setting the top mass=175 GeV within HitFit should favor the correct assignment category. This is proved in the following figure (figure 5) where we plot the number of times that the lowest χ^2 permutation gives the correct or wrong assignment.
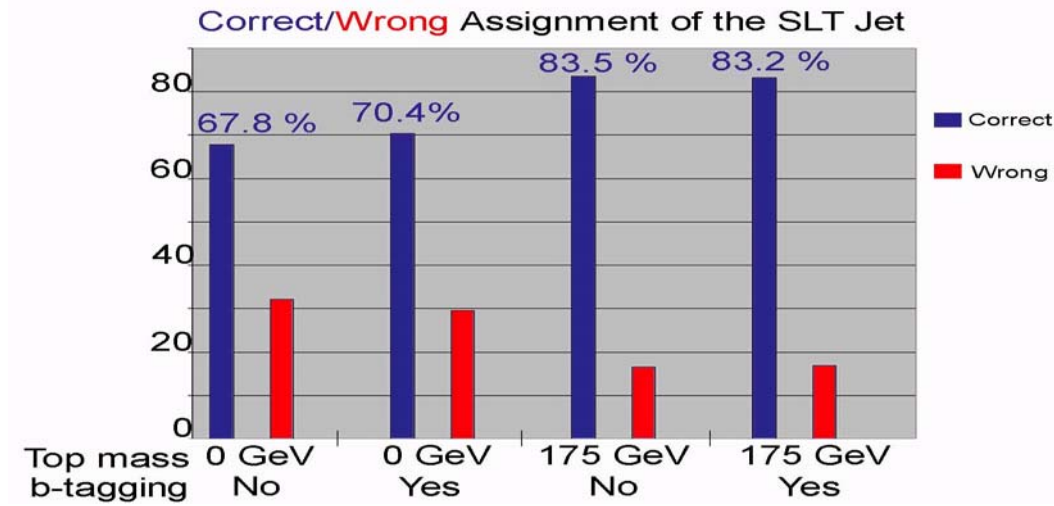

Figure 5. This figure shows the percentage of getting the correct/wrong assignment for different options within HitFit. Note that we only look at the permutation with the lowestχ^2 and the plotted events pass all the cuts of section II.2 without a P_T^{Rel} cut.

5. Combining the sign and assignment studies

At this point we will combine our studies for the sign between the daughter muon and the parent b quark with our studies for the assignment of the b to top quarks. So far we have discussed several methods for improving independently the sign and assignment techniques; however, we really want to study them together and find a quantity that must be optimized. By combining these studies we will end up with the following two categories:

- Correct sign AND kinematic assignment (Nc is the total number of events in this category)
- Wrong sign OR kinematic assignment (Nw is the total number of events in this category)

We also use the variables Sign and Assignment in order to decide in which category we should place each event. We set Sign=1 (Sign=-1) in the case where the daughter muon and the parent b quark have the same sign (opposite signs). Similarly, we set Assignment=1 (Assignment=-1) when the b quark is assigned to the correct (wrong) top quark (i.e. Assignment=1 when bLep is assigned to tLep). Furthermore, events with Assignment*Sign>0 contributes to the "Correct sign AND assignment" category where events with Assignment*Sign<0 add to the "Wrong sign OR assignment" category.

The quantity that we want to optimize is called significance and is defined to be (for a derivation of this quantity see Appendix A):

$$Significance = D*sqrt(\varepsilon) * sqrt(Ntotal)$$

-"D" is called the dilution variable and is defined to be D=2c-1
-"c" is the percentage of getting an event in the "Correct sign AND kinematic assignment" (c=Nc/(Nc+Nw)).
-"Ntotal" is the total number of events that pass all the cuts (II.2 and II.4.iv) but without any P_T^{Rel} cut.
-"ε" stands for ε =Nused/Ntotal, where Nused is the total number of events that passed all the pre-selection cuts (II.2) , have a SLT (II.3.ii) , χ^2 to be positive, parton-to-jet assignment to be successful and any given P_T^{Rel} cut.

Basically, we create the parent sample without any P_T^{Rel} cuts. Then we try different values for P_T^{Rel} and check how the significance is changing. First we do this analysis without using b-tagging within HitFit and then we do the same analysis but this time with the use of b-tagging within HitFit. Based on the results of these two analyses we choose the one that gives better results for the significance.

Lastly, we use four samples with different mass constrains within HitFit (top mass=0 (default), top mass=165GeV, 175GeV and 185GeV) and try to optimize D^2e (or significance square). The following figure (fig. 6) gives D^2e versus P_T^{Rel} cuts.
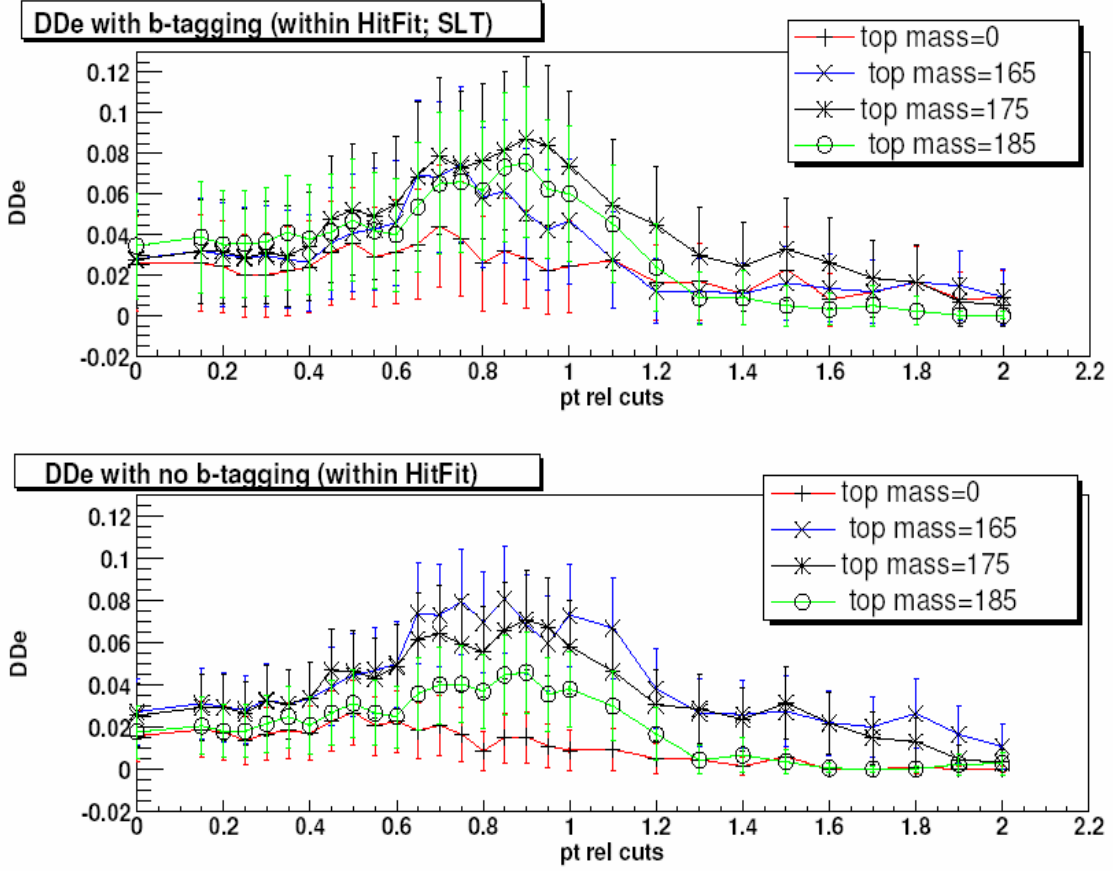
Figure 6. D^2e (unit less) versus P_T^{Rel} (GeV)cuts. The first and second plots are with SLT and no b-tagging respectively (within HitFit).

As we can see from fig.6, the use of b-tagging within HitFit helps us to get slightly better results for D^2e. On the other hand, constraining the top mass=175GeV inside HitFit gives a significant improvement on D^2e. These statements become more obvious if we take a look at fig.7.

Figure 7. Comparison of different techniques for the sample with top mass=175GeV.

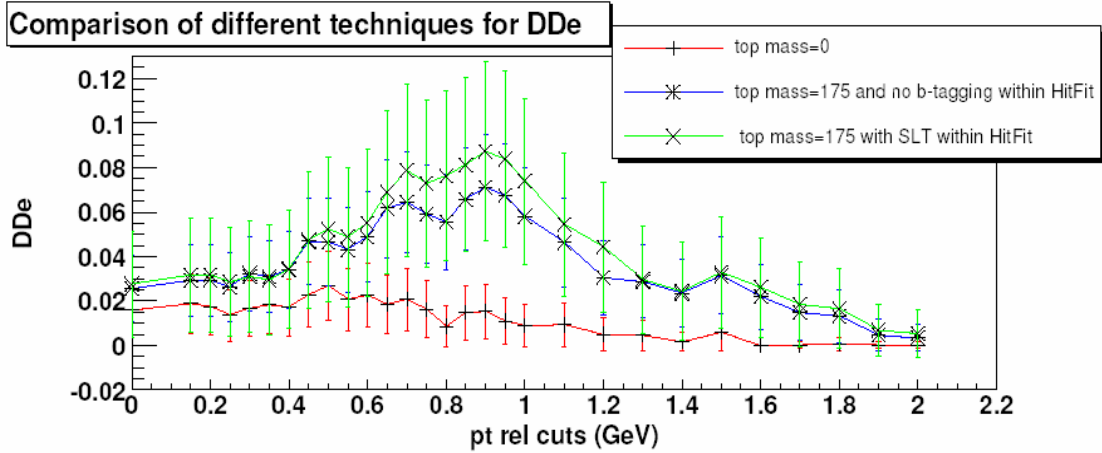As a final point, we want to draw attention to a technique that is used for the top mass analysis and might be useful for the charge measurement as well. This technique is called "Ideogram Method" and uses all the positive χ^2 solutions with a weighting factor. The complete formula of this method is given from the following expression [2]:

$$L(m_{top}) = \int dm' \ \Sigma w_i \ G(m', m_i, \sigma_I) \ [P*BW(m',m_{top}) + (1-P)*BG(m')]$$

-$w_i = \exp(-0.5*\chi^2)$.
-BW: Breit-Wigner normalized on whole interval m'.
-BG: Background shape from MC normalized on whole interval m'.

6. Calculating the total number of data events needed for a CL of 95%

In order to derive the significance (see section II.5) we assume that we have Poisson distribution for Nc and Nw because of the small number of events that pass the selection criteria for this analysis. However, a Poisson Statistics cannot characterize the difference of these two quantities.

The accurate way of calculating the total number of events is to perform a double integral over the two probabilities (Nc and Nw). On the other hand, we can ask for enough events to get a confidence level of 95% and then we should have enough events to get approximately Gaussian errors.

In the latter case the calculation of the total number of events is much simpler. With the use of a table of gaussian confidence level [3] and with the intention that we want at least a 95% confidence level (or a 5% one-sided confidence level) we get that the significance has to be greater than 1.64 sigma. We also know from section II.5 that the square of the significance peaks around 0.08 (for the sample with top mass=175 GeV and b-tagging within HitFit). Based on that we can calculate the total number of events:

$$\text{Significance} > 1.64 \rightarrow D^2*\varepsilon*N_{Total} > 2.69 \rightarrow 0.08*N_{Total} > 2.69$$

$$\mathbf{N_{Total} = 34 \pm 17 \ \ b\text{-tagged} \ \ events}$$

III. Signal and Background Studies

1. Introduction

After studying and understanding the tools that are used in this analysis, we continue by investigating how stable this analysis is on signal and background events. As mentioned earlier, we only use W+4jets (W+jjjj, W+cjjj, W+ccjj and Wbbjj) for background events. For signal events we use the same dataset as the one that we use for the signal studies; however, this time we treat it like data events ignoring information regarding generated particles (MC). The process and cuts are similar to the signal studies

but this time we calculate the percentage of getting the SM sign for the top quark. It is worth pointing out that if the analysis is stable, the top charge should be +2/3 for the reason that this analysis is performed on MC events.

2. Selection Requirements (Cuts)

The selection requirements are closely the same as in the signal studies. One can find the applied cuts in sections II.2., II.3.i and II.4.iv. . Obviously, the selection criteria that involve generated particles (MC) cannot be applied this time. Another important difference of this analysis and the signal studies is the assignment of bLep and bHad. In the signal studies we are able to distinguish the two b-quarks by using information from generated particles (MC); nonetheless, this is not the case for data events. Depending on which b-quark is b-tagged (bLep or bHad), the six permutations are different so we really have to decide which b-quark is b-tagged.

This ambiguity is solved with the use of the powerful HitFit package. As we can see from fig. 8, HitFit assigns the b-tagged jet correctly 83% of the cases. Based on this information we try both solutions and select the one that assigns the b-tagged jet correctly. To avoid any confusion, the statement "we try both solutions" denotes the fact that we first assume that the b-tagged jet is the bLep and we get the six permutations, which are applicable to this case (see section II.4.iv). Then we try the case where the b-tagged jet is the bHad and we get the six permutations that are relevant with b-tagging the bHad. Finally, we select the solution (bLep or bHad is b-tagged) where the lowest χ^2 gives the correct assignment for the b-tagged jet. In the case of two b-tagged jets, it does not matter which solution we choose since both solutions give the same answer. This technique does not always give the correct assignment but it gives the correct assignment more often than the wrong assignment.

We do this analysis for each sample separately and we record the sign of the top quark based on the kinematic assignment, the sign of the SLT muon and the sign of the isolated muon. There are eight possible combinations given that each quantity can take two values. From these combinations, four of them provide a sign of +2/3 and the other four give a sign of -4/3 for the top quark (see table 1).

Sign of the isolated μ	Sign of the SLT μ	Kinematic Assignment of the SLT jet	Top Charge
Positive	Positive	bLep	-4/3
Positive	Positive	bHad	+2/3
Positive	Negative	bLep	+2/3
Positive	Negative	bHad	-4/3
Negative	Positive	bLep	+2/3
Negative	Positive	bHad	-4/3
Negative	Negative	bLep	-4/3
Negative	Negative	bHad	+2/3

Table 1 . Determing the sign of the top quark.

3.Results

For the purposes of this analysis, we run topanalyze on each sample with different options inside HitFit. We create four different outputs for each sample with the following constrains within HitFit:

- Top mass=0 (default value) and SLT.
- Top mass=165 GeV and SLT.
- Top mass=175 GeV and SLT.
- Top mass=185 GeV and SLT.

We then run the code of this analusis and we record the results of each sample. After that, we continue by weighting them according to the cross section of each sample[2]. The cross sections (statistical error) for ttbar→μ+jets, W+jjjj, W+cjjj, W+ccjj and W+bbjj are 7.4pb (error) pb, 24.9 (error) pb, 1.75 (error) pb, 0.505 (error) pb and 0.663 (error) pb. In the next table we collect all the important numbers.

Top mass=0 (default value) and SLT within HitFit				
Sample	Number of events	Selected events	Percentage of getting the SM top charge	Weighted Contribution
ttbar		238	56.7 %	11.9 %
Wjjjj		3	66.7 %	47.2 %
Wcjjj		2	50 %	2.5 %
Wccjj		1	0 %	0 %
Wbbjj		15	66.7 %	1.3 %
Percentage of getting the SM top charge over all samples				**62.9 %**

Top mass=165GeV and SLT within HitFit				
Sample	Number of events	Selected events	Percentage of getting the SM top charge	Weighted Contribution
ttbar		235	56.6 %	11.8 %
Wjjjj		3	66.7 %	47.2 %
Wcjjj		2	50 %	2.5 %
Wccjj		1	100 %	1.4 %
Wbbjj		15	60 %	1.1 %
Percentage of getting the SM top charge over all samples				**64.0 %**

---

[2] The cross sections are taken from the DØnote 4141.

Top mass=175GeV and SLT within HitFit				
Sample	Number of events	Selected events	Percentage of getting the SM top charge	Weighted Contribution
ttbar		248	60.5 %	12.7 %
Wjjjj		3	66.7 %	47.2 %
Wcjjj		2	0 %	0 %
Wccjj		1	100 %	1.4 %
Wbbjj		15	53.3 %	1.0 %
Percentage of getting the SM top charge over all samples				**62.3 %**

Top mass=185GeV and SLT within HitFit				
Sample	Number of events	Selected events	Percentage of getting the SM top charge	Weighted Contribution
ttbar		252	65.1 %	13.7 %
Wjjjj		3	33.3 %	23.5 %
Wcjjj		2	50 %	2.5 %
Wccjj		1	100 %	1.4 %
Wbbjj		15	53.3 %	1.3 %
Percentage of getting the SM top charge over all samples				**42.4 %**

Table 2. Results of the signal and background studies.

IV. Conclusion

The analysis of the top charge determination, explained in this DØnote, shows that a measurement of the top charge is possible. The number of total b-tagged events allows us to hope for a first measurement of the top charge in the summer of 2004. However, table 2 indicates that a more careful analysis should be done.

One can see from table 2 that the contribution of W+jjjj is the biggest among all the other samples and only three events pass all the selection criteria. On the one hand, this is good and welcomed because we do not want a big number of background events. On the other hand, the low statistics do not allow us to investigate the behavior of the background events. Another interesting point is that this analysis is done on MC events with top charge of +2/3 and we think that the same analysis should be done for MC events with top charge -4/3. In conclusion, MC events should be generated for the purposes of the top charge studies.

Beside generating more MC events, one can also take advantage of techniques that are used from the top mass group. Two techniques that are worth trying are: i) use all permutation of the kinematic fitting and weight them according to their χ^2 (weighting factor= exp(-0.5*χ^2) and ii) in addition to SLT within SLT also use Secondary Vertex tagging (SVX) and Impact Parameter tagging. IP). The first technique should increase the

number of events and the second technique should reduce the background caused by incorrect assignment; remember, in the case of two b-tags we only get two permutations.

<u>Acknowledgments</u>

Appendix A

Here we derive the significance, which is defined in II.4 as the quantity that must be optimized for the signal studies.

$$\text{Significance} = (N_c - N_w)/\delta(N_c - N_w) = (N_c - N_w)/\text{sqrt}(\delta^2(N_c) + \delta^2(N_w))$$

We also know that $\delta(N_c) = \text{sqrt}(N_c)$ and $\delta(N_w) = \text{sqrt}(N_w)$ (Poison Statistics)

$$\Rightarrow \text{Significance} = (N_c - N_w)/\text{sqrt}(N_c + N_w)$$

We define the percentage of getting an event in the correct (c) and wrong (w) category with the following expressions:

$$c = N_c/(N_c + N_w) \quad \text{and} \quad w = N_w/(N_c + N_w)$$

We use that $N_{used} = N_c + N_w$ in order to simplify further the expression for the significance:

$$\text{Significance} = (N_c - N_w)/\text{sqrt}(N_c + N_w) = (cN_{used} - wN_{used})/\text{sqrt}(N_{used}) = (c - w)\text{sqrt}(N_{used})$$

It is obvious that $w + c = 1$ and we also define $e = N_{used}/N_{total}$, where N_{total} is the total number of events that passed all the pre-selection cuts before applying any P_μ^{Rel} cuts. In other words is the total number of events in the parent sample.

$$\text{Significance} = (2c-1)\text{sqrt}(e) * \text{sqrt}(N_{total})$$

Finally we use D to denote the dilution variable, which is equal to $D = 2c-1$.

$$\text{Significance} = D*\text{sqrt}(e)*\text{sqrt}(N_{total})$$

The quantity that we want to optimize is $D*\text{sqrt}(e)$ or $D^2 e$ since the quantity N_{total} is constant.

Reference:

[1]
[2]
[3]